RAW SEQUENCE LISTING   DATE: 12/28/2001
PATENT APPLICATION: US/09/424,458   TIME: 16:10:51

Input Set : A:\13198.asc
Output Set: N:\CRF3\12282001\I424458.raw

```
 4 <110> APPLICANT: Hayward, Nicholas
 5        Silins, Ginters
 6        Grimmond, Sean
 7        Gartside, Michael
 8        Hancock, John
10 <120> TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
12 <130> FILE REFERENCE: 13198
14 <140> CURRENT APPLICATION NUMBER: 09/424,458
15 <141> CURRENT FILING DATE: 2000-03-16
17 <150> PRIOR APPLICATION NUMBER: PCT/AU98/00380
18 <151> PRIOR FILING DATE: 1998-05-22
20 <160> NUMBER OF SEQ ID NOS: 111
22 <170> SOFTWARE: PatentIn Ver. 2.1
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 8
26 <212> TYPE: PRT
27 <213> ORGANISM: Artificial Sequence
29 <220> FEATURE:
30 <223> OTHER INFORMATION: Description of Artificial Sequence:Peptide repeat
31        motif in DnaJ homologues.
33 <220> FEATURE:
34 <221> NAME/KEY: UNSURE
35 <222> LOCATION: (2)..(3)
36 <223> OTHER INFORMATION: Xaa at position 2,3 can be any amino acid.
38 <220> FEATURE:
39 <221> NAME/KEY: UNSURE
40 <222> LOCATION: (5)
41 <223> OTHER INFORMATION: Xaa at position 5 can be any amino acid.
43 <220> FEATURE:
44 <221> NAME/KEY: UNSURE
45 <222> LOCATION: (7)
46 <223> OTHER INFORMATION: Xaa at position 7 can be any amio acid.
48 <400> SEQUENCE: 1
49 Cys Xaa Xaa Cys Xaa Gly Xaa Gly
50    1               5
53 <210> SEQ ID NO: 2
54 <211> LENGTH: 1242
55 <212> TYPE: DNA
56 <213> ORGANISM: Homo sapiens
58 <220> FEATURE:
59 <221> NAME/KEY: CDS
60 <222> LOCATION: (30)..(959)
62 <400> SEQUENCE: 2
63 tcagtaaaca cagagactgg ggatcgatc atg ggg ctt tgt aag tgc ccc aag    53
65                                 Met Gly Leu Cys Lys Cys Pro Lys
66                                  1               5
68 aga aag gtg acc aac ctg ttc tgc ttc gaa cat cgg gtc aac gtc tgc   101
```

```
70 Arg Lys Val Thr Asn Leu Phe Cys Phe Glu His Arg Val Asn Val Cys
71      10                  15                  20
73 gag cac tgc ctg gta gcc aat cac gcc aag tgc atc gtc cag tcc tac    149
75 Glu His Cys Leu Val Ala Asn His Ala Lys Cys Ile Val Gln Ser Tyr
76   25                  30                  35                  40
78 ctg caa tgg ctc caa gat agc gac tac aac ccc aat tgc cgc ctg tgc    197
80 Leu Gln Trp Leu Gln Asp Ser Asp Tyr Asn Pro Asn Cys Arg Leu Cys
81                  45                  50                  55
83 aac ata ccc ctg gcc agc cga gag acg acc cgc ctt gtc tgc tat gat    245
85 Asn Ile Pro Leu Ala Ser Arg Glu Thr Thr Arg Leu Val Cys Tyr Asp
86                  60                  65                  70
88 ctc ttt cac tgg gcc tgc ctc aat gaa cgt gct gcc cag cta ccc cga    293
90 Leu Phe His Trp Ala Cys Leu Asn Glu Arg Ala Ala Gln Leu Pro Arg
91              75          ·       80                  85
93 aac acg gca cct gcc ggc tat cag tgc ccc agc tgc aat ggc ccc atc    341
95 Asn Thr Ala Pro Ala Gly Tyr Gln Cys Pro Ser Cys Asn Gly Pro Ile
96         90                  95                  100
98 ttc ccc cca acc aac ctg gct ggc ccc gtg gcc tcc gca ctg aga gag    389
100 Phe Pro Pro Thr Asn Leu Ala Gly Pro Val Ala Ser Ala Leu Arg Glu
101 105                 110                 115                 120
103 aag ctg gcc aca gtc aac tgg gcc cgg gca gga ctg ggc ctc cct ctg    437
105 Lys Leu Ala Thr Val Asn Trp Ala Arg Ala Gly Leu Gly Leu Pro Leu
106                 125                 130                 135
108 atc gat gag gtg gtg agc cca gag ccc gag ccc ctc aac acg tct gac    485
110 Ile Asp Glu Val Val Ser Pro Glu Pro Glu Pro Leu Asn Thr Ser Asp
111                 140                 145                 150
113 ttc tct gac tgg tct agt ttt aat gcc agc agt acc cct gga cca gag    533
116 Phe Ser Asp Trp Ser Ser Phe Asn Ala Ser Ser Thr Pro Gly Pro Glu
117                 155                 160                 165
119 gag gta gac agc gcc tct gct gcc cca gcc ttc tac agc cga gcc ccc    581
121 Glu Val Asp Ser Ala Ser Ala Ala Pro Ala Phe Tyr Ser Arg Ala Pro
122         170                 175                 180
124 cgg ccc cca gct tcc cca ggc cgg ccc gag cag cac aca gtg atc cac    629
126 Arg Pro Pro Ala Ser Pro Gly Arg Pro Glu Gln His Thr Val Ile His
127 185                 190                 195                 200
129 atg ggc aat cct gag ccc ttg act cac gcc cct agg aag gtg tat gat    677
131 Met Gly Asn Pro Glu Pro Leu Thr His Ala Pro Arg Lys Val Tyr Asp
132                 205                 210                 215
134 acg cgg gat gat gac cgg aca cca ggc ctc cat gga gac tgt gac gat    725
136 Thr Arg Asp Asp Asp Arg Thr Pro Gly Leu His Gly Asp Cys Asp Asp
137                 220                 225                 230
139 gac aag tac cga cgt cgg ccg gcc ttg ggt tgg ctg gcc cgg ctg cta    773
141 Asp Lys Tyr Arg Arg Arg Pro Ala Leu Gly Trp Leu Ala Arg Leu Leu
142         235                 240                 245
144 agg agc cgg gct ggg tct cgg aag cgg ccg ctg acc ctg ctc cag cgg    821
146 Arg Ser Arg Ala Gly Ser Arg Lys Arg Pro Leu Thr Leu Leu Gln Arg
147         250                 255                 260
149 gcg ggg ctg ctg cta ctc ttg gga ctg ctg ggc ttc ctg gcc ctc ctt    869
151 Ala Gly Leu Leu Leu Leu Leu Gly Leu Leu Gly Phe Leu Ala Leu Leu
```

RAW SEQUENCE LISTING   DATE: 12/28/2001
PATENT APPLICATION: US/09/424,458   TIME: 16:10:51

Input Set : A:\13198.asc
Output Set: N:\CRF3\12282001\I424458.raw

```
152 265                  ._ 270    ._     _.  :      275                      280
154 gcc ctc atg tct cgc cta ggc cgg gcc gca gct gac agc gat ccc aac     917
156 Ala Leu Met Ser Arg Leu Gly Arg Ala Ala Ala Asp Ser Asp Pro Asn
157                 285                   290                 295
159 ctg gac cca ctc atg aac cct cac atc cgc gtg ggc ccc tcc             959
161 Leu Asp Pro Leu Met Asn Pro His Ile Arg Val Gly Pro Ser
162             300                   305                 310
164 tgagcccccct tgcttgtggc taggccagcc taggatgtgg gttctgtgga ggagaggcgg 1019
167 ggtaatgggg aggctgaggg cacctcttca ctgcccctct ccctcaagcc taagacacta 1079
170 agaccccaga cccaaagcca agtccaccag agtggctcgc aggccaggcc tggagtcccc 1139
173 gtgggtcaag catttgtctt gacttgcttt ctcccgggtc tccagcctcc gacccctcgc 1199
176 cccatgaagg agctggcagg tggaaataaa caacaacttt att                     1242
180 <210> SEQ ID NO: 3
181 <211> LENGTH: 310
182 <212> TYPE: PRT
183 <213> ORGANISM: Homo sapiens
185 <400> SEQUENCE: 3
186 Met Gly Leu Cys Lys Cys Pro Lys Arg Lys Val Thr Asn Leu Phe Cys
187  1               5                  10                  15
189 Phe Glu His Arg Val Asn Val Cys Glu His Cys Leu Val Ala Asn His
190             20                  25                  30
192 Ala Lys Cys Ile Val Gln Ser Tyr Leu Gln Trp Leu Gln Asp Ser Asp
193         35                  40                  45
195 Tyr Asn Pro Asn Cys Arg Leu Cys Asn Ile Pro Leu Ala Ser Arg Glu
196     50                  55                  60
198 Thr Thr Arg Leu Val Cys Tyr Asp Leu Phe His Trp Ala Cys Leu Asn
199 65                  70                  75                  80
201 Glu Arg Ala Ala Gln Leu Pro Arg Asn Thr Ala Pro Ala Gly Tyr Gln
202                 85                  90                  95
204 Cys Pro Ser Cys Asn Gly Pro Ile Phe Pro Pro Thr Asn Leu Ala Gly
205             100                 105                 110
207 Pro Val Ala Ser Ala Leu Arg Glu Lys Leu Ala Thr Val Asn Trp Ala
208         115                 120                 125
210 Arg Ala Gly Leu Gly Leu Pro Leu Ile Asp Glu Val Val Ser Pro Glu
211     130                 135                 140
213 Pro Glu Pro Leu Asn Thr Ser Asp Phe Ser Asp Trp Ser Ser Phe Asn
214 145                 150                 155                 160
216 Ala Ser Ser Thr Pro Gly Pro Glu Glu Val Asp Ser Ala Ser Ala Ala
217             165                 170                 175
219 Pro Ala Phe Tyr Ser Arg Ala Pro Arg Pro Pro Ala Ser Pro Gly Arg
220             180                 185                 190
222 Pro Glu Gln His Thr Val Ile His Met Gly Asn Pro Glu Pro Leu Thr
223         195                 200                 205
226 His Ala Pro Arg Lys Val Tyr Asp Thr Arg Asp Asp Asp Arg Thr Pro
227     210                 215                 220
230 Gly Leu His Gly Asp Cys Asp Asp Asp Lys Tyr Arg Arg Arg Pro Ala
231 225                 230                 235                 240
233 Leu Gly Trp Leu Ala Arg Leu Leu Arg Ser Arg Ala Gly Ser Arg Lys
234             245                 250                 255
```

RAW SEQUENCE LISTING   DATE: 12/28/2001
PATENT APPLICATION: US/09/424,458   TIME: 16:10:51

Input Set : A:\13198.asc
Output Set: N:\CRF3\12282001\I424458.raw

```
236 Arg Pro Leu Thr Leu Leu Gln Arg Ala Gly Leu Leu Leu Leu Gly
237             260                   265               270
239 Leu Leu Gly Phe Leu Ala Leu Leu Ala Leu Met Ser Arg Leu Gly Arg
240           275                 280                 285
242 Ala Ala Ala Asp Ser Asp Pro Asn Leu Asp Pro Leu Met Asn Pro His
243       290                 295                 300
245 Ile Arg Val Gly Pro Ser
246 305                 310
251 <210> SEQ ID NO: 4
252 <211> LENGTH: 2415
253 <212> TYPE: DNA
254 <213> ORGANISM: Homo sapiens
256 <220> FEATURE:
257 <221> NAME/KEY: CDS
258 <222> LOCATION: (3)..(2186)
260 <400> SEQUENCE: 4
261 cg att tca ttc ctc gct ccc cac agg tcc ctc tcc cca aaa tat tcc   47
263    Ile Ser Phe Leu Ala Pro His Arg Ser Leu Ser Pro Lys Tyr Ser
264    1               5                  10                  15
266 cat ctt gtc cta gcc cat ccc cca gac tat ctc aag gac cag ctg tcc   95
268 His Leu Val Leu Ala His Pro Pro Asp Tyr Leu Lys Asp Gln Leu Ser
269                 20                  25                  30
271 cca cgc ccc cga cct cca cta ggc ctg tgc cac ccg ctg cct gca gga  143
273 Pro Arg Pro Arg Pro Pro Leu Gly Leu Cys His Pro Leu Pro Ala Gly
274             35                  40                  45
276 aga cgc ccg gtc ccg ggc cgg gtt agc ccc atg gga acg cag cgc ctg  191
278 Arg Arg Pro Val Pro Gly Arg Val Ser Pro Met Gly Thr Gln Arg Leu
279         50                  55                  60
281 tgt ggc cgc ggg act caa ggc tgg cct ggc tca agt gaa cag cac gtc  239
283 Cys Gly Arg Gly Thr Gln Gly Trp Pro Gly Ser Ser Glu Gln His Val
284     65                  70                  75
286 cag gag gcg acc tcg tcc gcg ggt ttg cat tct ggg gtg gac gag ctg  287
288 Gln Glu Ala Thr Ser Ser Ala Gly Leu His Ser Gly Val Asp Glu Leu
289 80                  85                  90                  95
291 ggg gtt cgg tcc gag ccc ggt ggg agg ctc ccg gag cgc agc ctg ggc  335
293 Gly Val Arg Ser Glu Pro Gly Gly Arg Leu Pro Glu Arg Ser Leu Gly
294                 100                 105                 110
296 cca gcc cac ccc gcg ccg gcg gcc atg gca ggc acc ctg gac ctg gac  383
298 Pro Ala His Pro Ala Pro Ala Ala Met Ala Gly Thr Leu Asp Leu Asp
299             115                 120                 125
301 aag ggc tgc acg gtg gag gag ctg ctc cgc ggg tgc atc gaa gcc ttc  431
303 Lys Gly Cys Thr Val Glu Glu Leu Leu Arg Gly Cys Ile Glu Ala Phe
304         130                 135                 140
306 gat gac tcc ggg aag gtg cgg gac ccg cag ctg gtg cgc atg ttc ctc  479
308 Asp Asp Ser Gly Lys Val Arg Asp Pro Gln Leu Val Arg Met Phe Leu
309     145                 150                 155
311 atg atg cac ccc tgg tac atc ccc tcc tct cag ctg gcg gcc aag ctg  527
313 Met Met His Pro Trp Tyr Ile Pro Ser Ser Gln Leu Ala Ala Lys Leu
314 160                 165                 170                 175
```

```
316 ctc cac atc tac caa caa tcc cgg aag gac aac tcc aat tcc ctg cag   575
318 Leu His Ile Tyr Gln Gln Ser Arg Lys Asp Asn Ser Asn Ser Leu Gln
319                 180             185                 190
321 gtg aaa acg tgc cac ctg gtc agg tac tgg atc tcc gcc ttc cca gcg   623
323 Val Lys Thr Cys His Leu Val Arg Tyr Trp Ile Ser Ala Phe Pro Ala
324             195                 200                 205
326 gag ttt gac ttg aac ccg gag ttg gct gag cag atc aag gag ctg aag   671
328 Glu Phe Asp Leu Asn Pro Glu Leu Ala Glu Gln Ile Lys Glu Leu Lys
329             210                 215                 220
331 gct ctg cta gac caa gaa ggg aac cga cgg cac agc agc cta atc gac   719
333 Ala Leu Leu Asp Gln Glu Gly Asn Arg Arg His Ser Ser Leu Ile Asp
334         225                 230                 235
336 ata gac agc gtc cct acc tac aag tgg aag cgg cag gtg act cag cgg   767
338 Ile Asp Ser Val Pro Thr Tyr Lys Trp Lys Arg Gln Val Thr Gln Arg
339 240                 245                 250                 255
341 aac cct gtg gga cag aaa aag cgc aag atg tcc ctg ttg ttt gac cac   815
343 Asn Pro Val Gly Gln Lys Lys Arg Lys Met Ser Leu Leu Phe Asp His
344                 260                 265                 270
346 ctg gag ccc atg gag ctg gcg gag cat ctc acc tac ttg gag tat cgc   863
348 Leu Glu Pro Met Glu Leu Ala Glu His Leu Thr Tyr Leu Glu Tyr Arg
349             275                 280                 285
351 tcc ttc tgc aag atc ctg ttt cag gac tat cac agt ttc gtg act cat   911
353 Ser Phe Cys Lys Ile Leu Phe Gln Asp Tyr His Ser Phe Val Thr His
354             290                 295                 300
356 ggc tgc act gtg gac aac ccc gtc ctg gag cgg ttc atc tcc ctc ttc   959
358 Gly Cys Thr Val Asp Asn Pro Val Leu Glu Arg Phe Ile Ser Leu Phe
359         305                 310                 315
361 aac agc gtc tca cag tgg gtg cag ctc atg atc ctc agc aaa ccc aca   1007
363 Asn Ser Val Ser Gln Trp Val Gln Leu Met Ile Leu Ser Lys Pro Thr
364 320                 325                 330                 335
366 gcc ccg cag cgg gcc ctg gtc atc aca cac ttt gtc cac gtg gcg gag   1055
368 Ala Pro Gln Arg Ala Leu Val Ile Thr His Phe Val His Val Ala Glu
369             340                 345                 350
371 aag ctg cta cag ctg cag aac ttc aac acg ctg atg gca gtg gtc ggg   1103
373 Lys Leu Leu Gln Leu Gln Asn Phe Asn Thr Leu Met Ala Val Val Gly
374         355                 360                 365
376 ggc ctg agc cac agc tcc atc tcc cgc ctc aag gag acc cac agc cac   1151
378 Gly Leu Ser His Ser Ser Ile Ser Arg Leu Lys Glu Thr His Ser His
379         370                 375                 380
381 gtt agc cct gag acc atc aag ctc tgg gag ggt ctc acg gaa cta gtg   1199
383 Val Ser Pro Glu Thr Ile Lys Leu Trp Glu Gly Leu Thr Glu Leu Val
384         385                 390                 395
386 acg gcg aca ggc aac tat ggc aac tac cgg cgt cgg ctg gca gcc tgt   1247
388 Thr Ala Thr Gly Asn Tyr Gly Asn Tyr Arg Arg Arg Leu Ala Ala Cys
389 400                 405                 410                 415
391 gtg ggc ttc cgc ttc ccg atc ctg ggt gtg cac ctc aag gac ctg gtg   1295
393 Val Gly Phe Arg Phe Pro Ile Leu Gly Val His Leu Lys Asp Leu Val
394                 420                 425                 430
396 gcc ctg cag ctg gca ctg cct gac tgg ctg gac cca gcc cgg acc cgg   1343
```

VERIFICATION SUMMARY   DATE: 12/28/2001
PATENT APPLICATION: US/09/424,458   TIME: 16:10:52

Input Set : **A:\13198.asc**
Output Set: **N:\CRF3\12282001\I424458.raw**

```
L:49 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:1758 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49
L:1761 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49
L:1767 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49
L:1770 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49
L:1773 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49
L:2163 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74
L:2166 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74
L:2169 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74
L:2172 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74
L:2175 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74
L:2178 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74
L:2212 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:75
L:2215 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:75
L:2254 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:76
L:2257 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:76
```